An Empirical Process Central Limit Theorem for Multidimensional Dependent Data*

Olivier Durieu[†], Marco Tusche[‡]

October 1., 2012

Abstract Let $(U_n(t))_{t\in\mathbb{R}^d}$ be the empirical process associated to an $\mathbb{R}^d$-valued stationary process $(X_i)_{i\geq 0}$. In the present paper, we introduce very general conditions for weak convergence of $(U_n(t))_{t\in\mathbb{R}^d}$, which only involve properties of processes $(f(X_i))_{i\geq 0}$ for a restricted class of functions $f\in\mathcal{G}$. Our results significantly improve those of [8] and [7] and provide new applications.

The central interest in our approach is, that it does not need the indicator functions, which define the empirical process $(U_n(t))_{t\in\mathbb{R}^d}$, itself to belong to the class $\mathcal{G}$. This is particularly useful when dealing with data arising from dynamical systems or functionals of Markov chains. In the proofs we make use of a new application of a chaining argument and generalize ideas first introduced in [8] and [7].

Finally we will show how our general conditions apply in the case of multiple mixing processes of polynomial decrease and causal functions of independent and identically distributed processes, which could not be treated by the preceding results in [8] and [7].

1 Introduction

The present paper concerns the question of the weak convergence of empirical processes under weak dependence of the underlying process. Let us consider some stationary $\mathbb{R}^d$-valued process[1]

---

*The final publication is available at
http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s10959-012-0450-3

†Laboratoire de Mathématiques et Physique Théorique (UMR CNRS 7350), Fédération Denis Poisson (FR CNRS 2964), Université de Tours, Parc de Grandmont, 37200 Tours, France, olivier.durieu@lmpt.univ-tours.fr, Fax: (+33) 247 367 068, Fon: (+33) 247 367 421

‡Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstraße 150, 44780 Bochum, Germany, marco.tusche@RUB.de, Fax: (+49) 234 32 14039, Fon: (+49)234 32 23425

[1]$\mathbb{N} := \{0, 1, \ldots\}$, $\mathbb{N}^* := \mathbb{N}\setminus\{0\}$

$(X_i)_{i\in\mathbb{N}}$. The empirical process of $(X_i)_{i\in\mathbb{N}}$ is the process $(U_n)_{n\in\mathbb{N}^*}$ of $\mathbb{D}([-\infty,\infty]^d)$-valued[2] random variables

$$U_n := (U_n(t_1,\ldots,t_d))_{(t_1,\ldots,t_d)\in[-\infty,\infty]^d}$$

given by

$$U_n(t_1,\ldots,t_d) := \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\mathbf{1}_{\prod_{j=1}^{d}[-\infty,t_j]}(X_i) - \mathrm{E}\mathbf{1}_{\prod_{j=1}^{d}[-\infty,t_j]}(X_0)\Big).$$

The study of the weak convergence of such a process began with Donsker [10], who proved convergence of the empirical process to some Gaussian process in the case of i.i.d. $\mathbb{R}$-valued data. Donsker's result has been extended to sequences of weakly dependent ($\mathbb{R}$ or $\mathbb{R}^d$-valued) random variables by many authors. Among others, it shall be remarked that Billingsley [2] gave a result for functionals of φ-mixing sequences, Berkes and Philipp [1] under strong mixing assumptions, Doukhan, Massart and Rio [12] for absolutely regular sequences, Borovkova, Burton and Dehling [3] for functionals of absolutely regular processes, Doukhan and Louhichi [11], and Dedecker and Prieur [6] for new dependence coefficients.

The technique of the proof usually consists of establishing finite-dimensional convergence of the process and to use an appropriate moment bound to get tightness of the process (see [9]).

In the recent paper [8], the authors introduced a new approach which is useful when the required properties on the underlying process are only established for a class of functions not containing the indicators. The required conditions are the Central Limit Theorem and a bound of the 4-th moment of partial sums for the process $(f(X_i))_{i\geq0}$, when f belongs to the class of Lipschitz functions. Under these conditions, in the case of $\mathbb{R}$-valued data, they were able to prove an empirical CLT. Later, in [7], the technique was adapted to treat the case of $\mathbb{R}^d$-valued data and using the class of Hölder functions. In that case bounds on higher moments are needed.

These required conditions can be satisfied in cases where the data arise from a dynamical system. In particular, for a large class of dynamical systems, they can be derived from the study of the so called Perron-Frobenius operator. More generally, they can be established in cases where the system presents some multiple mixing properties (see Section 2 and [7] for more details).

The aim of the present paper is to give a very general setting under which the technique of [8] and [7] can be applied. In this way, we obtain extensions of the previous results. In particular, we show that the approach of [8] and [7] can be applied with more general classes of functions than Lipschitz and Hölder functions. We provide general conditions on the class of functions for which the CLT and moment bounds have to hold. In addition, we use a more general moment bound, balancing conditions on the distribution of X_0 with properties of the class of functions.

Our results allow us to handle processes with weaker moment bound than in previous works. In particular, we are able to treat systems satisfying a multiple mixing property with polynomial rates, whereas in [7] exponential rates were required. As a concrete application, we obtain new result for causal functions of i.i.d. processes.

---

[2]Let $\mathbb{D}([-\infty,\infty]^d)$ be the space of generalized multidimensional càdlàg functions $[-\infty,\infty]^d \to \mathbb{R}$ (for definition see [16, p.1286]), equipped with the multidimensional Skorohod metric d_0 as introduced in [16, p.1289] (see also [17]). Note that $(\mathbb{D}([-\infty,\infty]^d), d_0)$ is a complete and separable space (more precisely, [16] and [17] proved this for the space $\mathbb{D}([0,1]^d)$, but – since $[0,1]$ and $[-\infty,\infty]$ are homeomorphic – the metric on $\mathbb{D}([-\infty,\infty]^d)$ can naturally be extended to a metric on $\mathbb{D}([-\infty,\infty]^d)$ which conserves all relevant properties (c.f. [7, p.1081f.])).

2 Central Conditions and Statement of Main Results

2.1 Conditions

Before coming to the statement of our main results, let us have a look at our conditions, what they are needed for and some example of how they can be established.

As pointed out in the introduction, the present paper presents a technique which differs from the usual finite-dimensional convergence plus tightness of U_n approach. This is useful in cases, where one cannot or not without great effort directly prove these conditions, but where one can establish similar results for functions $(f(X_i))_{i \in \mathbb{N}}$ of the underlying process. These situations appear, e.g. when dealing with data arising from Markov chains or dynamical systems (c.f. [8] and [7]).

2.1.1 Central Limit Theorem under a Class of Functions.

We say that a process $(X_i)_{i \in \mathbb{N}}$ satisfies a Central Limit Theorem under a class of functions $\mathcal{G}$, if for every $f \in \mathcal{G}$ such that $\mathrm{E}(f(X_0)) = 0$, there exists $\sigma_f^2 < \infty$ such that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} f(X_i) \xrightarrow{\mathcal{D}} N(0, \sigma_f^2). \tag{1}$$

A lot of research has been devoted to establish CLTs under various classes of functions, such as functions of bounded variation, Lipschitz functions and Hölder functions. Hennion and Hervé [13] give a survey of such results for dynamical systems and Markov chains, where the Perron-Frobenius operator or the Markov operator satisfies a spectral gap condition. Another example are ergodic torus automorphisms, for which Leonov [15] and Le Borgne [14] established CLTs for Hölder functions. In the present paper, we consider causal functions of i.i.d. processes where the CLT holds for Hölder functions. CLTs for Lipschitz functions were studied for random iterative Lipschitz models and linear function of i.i.d. processes in [8].

2.1.2 Moment Bounds under a Class of Functions

To prove tightness, one usually makes use of conditions on the moments of increments of the empirical process. In order to restrict our conditions to functions of the process we will work with the following type of moment bounds under the same function space $\mathcal{G}$ for which the CLT (1) holds:

There are finite constants $C > 0$, $r \geq 1$, $p \in \mathbb{N}^*$ and nondecreasing functions $\Phi_1, \ldots, \Phi_p : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ such that for all $f \in \mathcal{G}$ with $\|f\|_\infty \leq 1$ and all $n \in \mathbb{N}^*$ we have

$$\mathrm{E}\left(\left| \sum_{i=1}^{n} f(X_i) - \mathrm{E}(f(X_0)) \right|^{2p} \right)$$

$$\leq C \sum_{i=1}^{p} n^i \, \|f(X_0) - \mathrm{E}(f(X_0))\|_r^i \, \Phi_i(\|f - \mathrm{E}(f(X_0))\|_\mathcal{G}). \tag{2}$$

This condition is met, e.g. for processes satisfying a multiple mixing property.

Definition 1 (Multiple mixing property)**.** Let $(X_i)_{i \in \mathbb{N}}$ be a stationary stochastic process of $\mathbb{R}^d$-valued random variables and let $\mathcal{G}$ be a class of measurable real-valued functions defined on $\mathbb{R}^d$ and equipped with a seminorm $\|\cdot\|_\mathcal{G}$. For integers $i_1, ..., i_j$, we write $i_j^* := i_1 + \ldots + i_j$. We say that $(X_n)_{n \in \mathbb{N}}$ is Θ, r-multiple mixing with respect to $\mathcal{G}$ if there exist a constant $r \in [1, \infty)$ and

a nonincreasing function $\Theta : \mathbb{N} \longrightarrow \mathbb{R}_0^+$ such that for any $p \in \mathbb{N}^*$ there is a constant $K_p < \infty$ satisfying:

$$
\begin{aligned}
&\left| \mathrm{Cov}\left(f(X_0)f(X_{i_1^*}) \cdot \ldots \cdot f(X_{i_{q-1}^*}), \, f(X_{i_q^*})f(X_{i_{q+1}^*}) \cdot \ldots \cdot f(X_{i_p^*}) \right) \right| \\
&\leq K_p \|f(X_0)\|_r \|f\|_{\mathcal{G}} \Theta(i_q)
\end{aligned}
\tag{3}
$$

for all $f \in \mathcal{G}$ with $\|f\|_\infty \leq 1$ and $\mathrm{E}(f(X_0)) = 0$ all $i_1, \ldots, i_p \in \mathbb{N}$, $q \in \{1, \ldots, p\}$. In that case we write $(X_n)_{n \in \mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{G})$.

The following proposition shows that multiple mixing systems satisfy a $2p$-th moment bound (2).

Proposition 1. *Let* $(X_n)_{n \in \mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{G})$ *for a function* Θ *such that there exists* $p \in \mathbb{N}^*$ *satisfying*

$$
\sum_{i=0}^{\infty} i^{2p-2} \Theta(i) < \infty.
\tag{4}
$$

Then there is a $C > 0$ *such that* (2) *holds for all* $f \in \mathcal{G}$ *such that* $\|f\|_\infty \leq 1$, *with* r *and* p *as above and* $\Phi_i(x) = x^i$.

The proof of Proposition 1 will be given in Section 5.

Note that similar moment bounds have been obtained in [7] under the stronger assumption of multiple mixing with exponential rate, i.e. when $\Theta(i)$ decreases exponentially. Under this assumption, one can get sharper moment bounds, with $\Phi_i(x) = \log^{2p-i}(x+1)$. The moment bounds obtained in the present paper are sufficient to apply Theorem 1, and thus our generalization allows one to treat processes that are multiple mixing with polynomial rate.

2.1.3 Approximation of the Indicator Functions

Conditions (1) and (2) refer to the processes $(f(X_i))_{i \in \mathbb{N}}$ for $f \in \mathcal{G}$. In order to obtain results for the empirical process U_n, we need to approximate the indicator functions occurring in the definition of U_n by functions from $\mathcal{G}$. To describe the quality of this approximation with respect to $\| \cdot \|_{\mathcal{G}}$, we introduce the upcoming definition.

In the following, $\leq$, $<$, ... used in $\mathbb{R}^d$ are to be understood component-wise[3] and sets such as $\{x \in [-\infty, \infty]^d : a \leq x < b\}$ are denoted by $(a, b]$. We will also use the notation w_g for the modulus of continuity of a real-valued function g, which is defined by

$$
w_g(\delta) := \sup\{|g(t) - g(s)| : s, t \in \mathbb{R}^d, \ \|t - s\| \leq \delta\}.
$$

Definition 2 ($\mathcal{G}$-F-control function). Let $\mathcal{G}$ be some vector space of real-valued functions defined on $\mathbb{R}^d$, equipped with a seminorm $\| \cdot \|_{\mathcal{G}}$, and let F denote the multidimensional distribution function of an $\mathbb{R}^d$-valued random variable X. We call a nondecreasing function $\Psi : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ a $\mathcal{G}$-F-control, if for every $a < b \in [-\infty, \infty]^d$ there is a function $\varphi_{(a,b)} \in \mathcal{G}$ such that for any $x \in \mathbb{R}^d$

$$
\mathbf{1}_{(-\infty, a]} \leq \varphi_{(a,b)} \leq \mathbf{1}_{(-\infty, b]}.
\tag{5}
$$

and such that

$$
\|\varphi_{(a,b)}\|_{\mathcal{G}} \leq \Psi\left(\frac{1}{\min\limits_{i=1,\ldots,d} w_{F_i}(b_i - a_i)} \right),
\tag{6}
$$

---

[3]i.e. for $a = (a_1, \ldots, a_d)$, $b = (b_1, \ldots, b_d) \in [-\infty, \infty]^d$ write $a \leq b$ if and only if $a_i \leq b_i$ for all $i = 1, \ldots, d$

where F_i denotes the i-th marginal distribution function of X.

If such a function Ψ exists, we say that $\mathcal{G}$ approximates the indicator functions (of rectangles $[-\infty, t]$, $t \in \mathbb{R}^d$) with $\mathcal{G}$-F-control Ψ.

Example 1. As an example consider the class of bounded α-Hölder functions $\mathcal{H}_\alpha$ equipped with the α-Hölder norm

$$\|f\|_{\mathcal{H}_\alpha} := \|f\|_\infty + \sup\left\{\frac{|f(x) - f(y)|}{\|x - y\|^\alpha} : x, y \in \mathbb{R}^d, \ x \neq y\right\}.$$

Choose $\varphi_{(a,b)} \in \mathcal{H}_\alpha$ as

$$\varphi_{(a,b)}(x_1, \ldots, x_d) := \prod_{i=1}^d \varphi\left(\mathbf{1}_{(-\infty,\infty)^2}(a_i, b_i) \cdot \frac{x_i - b_i}{b_i - a_i}\right),$$

where $\varphi : \mathbb{R} \to \mathbb{R}$ is given by $\varphi(x) = \mathbf{1}_{[-\infty,-1]}(x) - x\mathbf{1}_{(-1,0]}(x)$. Obviously this choice of $\varphi_{(a,b)}$ satisfies (5). Let us now check condition (6). Since for all $j = 1, \ldots, d$, $w_{F_j}(\delta) \leq w_F(\delta)$, we have

$$b_j - a_j \geq \inf\left\{\delta > 0 : w_F(\delta) \geq \min_{i=1,\ldots,d} w_{F_i}(b_i - a_i)\right\}.$$

Thus, by the definition of $\varphi_{(a,b)}$ we obtain

$$\|\varphi_{(a,b)}\|_{\mathcal{H}_\alpha} \leq d \max_{j=1,\ldots,d} \mathbf{1}_{(-\infty,\infty)^2}(a_j, b_j) \cdot \frac{1}{(b_j - a_j)^\alpha} + 1$$
$$\leq d\left(w_F^\leftarrow\left(\min_{i=1,\ldots,d} w_{F_i}(b_i - a_i)\right)\right)^{-\alpha} + 1,$$

where $w_F^\leftarrow(y) := \inf\left\{\delta > 0 : w_F(\delta) \geq y\right\}$. Hence (6) is satisfied for the nondecreasing function Ψ given by

$$\Psi(z) := d\left(w_F^\leftarrow\left(\frac{1}{z}\right)\right)^{-\alpha} + 1 \tag{7}$$

and thus Ψ defines an $\mathcal{H}_\alpha$-F-control, which gives us the following lemma:

Lemma 1. *The space of bounded α-Hölder functions $\mathcal{H}_\alpha$ approximate the indicator functions with $\mathcal{H}_\alpha$-F-control Ψ given by (7).*

2.2 Main Theorems

Our main result is the following:

Theorem 1. *Let $(X_i)_{i\in\mathbb{N}}$ be a stationary process of $\mathbb{R}^d$-valued random vectors with continuous multidimensional distribution function F. Assume that there is a vector space $\mathcal{G}$ of measurable functions $\mathbb{R}^d \to \mathbb{R}$, containing the constant functions, equipped with a seminorm $\|\cdot\|_\mathcal{G}$, satisfying the following conditions:*

  (i) For every $f \in \mathcal{G}$ such that $\|f\|_\infty < \infty$, the CLT (1) holds.

  (ii) $\mathcal{G}$ approximates the indicator functions with $\mathcal{G}$-F-control Ψ.

(iii) There are constants $r \geq 1$, $p > rd$, $\gamma_1, \ldots, \gamma_p$ satisfying

$$0 \leq \gamma_i < \frac{i}{r} + 2(p - i) - d \quad \text{for all } i = 1, \ldots, p, \tag{8}$$

and some nondecreasing functions $\Phi_1, \ldots, \Phi_p : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ satisfying

$$\Phi_i(2\Psi(z)) = \mathcal{O}(z^{\gamma_i}) \quad \text{as } z \to \infty, \tag{9}$$

such that for every $f \in \mathcal{G}$ with $\|f\|_\infty \leq 1$ the moment bound (2) holds.

Then there is a centered Gaussian process $(W(t))_{t \in [-\infty, \infty]^d}$ with almost surely continuous sample paths such that $U_n \xrightarrow{\mathcal{D}} W$, in the space $\mathbb{D}([-\infty, \infty]^d)$.

Remark 1. In fact, the functions $\Phi_1, \ldots, \Phi_p$ and the function Ψ only have to be nondecreasing for sufficiently large arguments. Note that condition (2) has only to be satisfied for a certain subclass of $\mathcal{G}$, see Remark 5.

As a consequence of this abstract theorem, we can give a statement for multiple mixing processes, for which conditions are more easily verifiable.

Theorem 2. *Let $(X_i)_{i \in \mathbb{N}}$ be a stationary $\mathbb{R}^d$-valued process with continuous multidimensional distribution function F. Assume there is a vector space $\mathcal{G}$ of measurable functions $\mathbb{R}^d \to \mathbb{R}$, containing the constant functions, equipped with a seminorm $\|\cdot\|_\mathcal{G}$ that satisfies the following conditions:*

(i) For every $f \in \mathcal{G}$ such that $\|f\|_\infty < \infty$ the CLT (1) holds.

(ii) The process $(X_i)_{i \in \mathbb{N}^*}$ is Θ, r-multiple mixing with respect to $\mathcal{G}$ for some $r \geq 1$ and a $\Theta : \mathbb{N} \to \mathbb{R}_0^+$, such that there exists a $p > dr$ satisfying $\sum_{i=0}^\infty i^{2p-2}\Theta(i) < \infty$.

(iii) $\mathcal{G}$ approximates the indicator functions with $\mathcal{G}$-F-control Ψ such that $\Psi(z) = \mathcal{O}(z^{\frac{1}{\gamma}})$ for some $\gamma > \frac{rp}{p-rd}$.

Then there is a centered Gaussian process $(W(t))_{t \in [-\infty, \infty]^d}$ with almost surely continuous sample paths such that $U_n \xrightarrow{\mathcal{D}} W$, in the space $\mathbb{D}([-\infty, \infty]^d)$.

Proof. By Proposition 1, (2) holds with $\Phi_i(x) = x^i$. Then, taking $\gamma_i = \frac{i}{\gamma}$, the assumptions of Theorem 1 are satisfied. $\qquad\square$

The rest of the paper is organized as follows: In Section 3 we present some particular cases of the theorem and some applications. The proof of Theorem 1 is given in Section 4. Section 5 is devoted to the proofs of the results of Section 3.

3 Applications

While Theorem 1 is stated in a very general context, this section is devoted to the special case where $\mathcal{G} = \mathcal{H}_\alpha$, the class of bounded α-Hölder functions. We will first obtain a corollary to Theorem 1 specifying the conditions for this choice of $\mathcal{G}$. As an application, we will study empirical processes of causal functions of an i.i.d. process.

3.1 Assumptions on Hölder continuous functions

The case of Hölder continuous functions has already been considered in [7]. In the present paper, we obtain generalizations of the results in [7], by allowing a larger class of functions Φ_i in the moment bound (2). In this way, we are able to treat processes that satisfy a multiple mixing condition with polynomial decay, rather than the exponential decay considered in [7].

Theorem 3. *Let $(X_i)_{i \in \mathbb{N}}$ be a stationary $\mathbb{R}^d$-valued process with continuous multidimensional distribution function F. Assume that:*

(i) *For every $f \in \mathcal{H}_\alpha$ the CLT (1) holds.*

(ii) *There are constants $\beta > 0$, $\gamma > 1$, $r \geq 1$, $p > rd$, $\kappa_1, \ldots, \kappa_p > 0$, $\lambda > 1$, $z_0 \in \mathbb{R}^+$ and an invertible increasing function $\Phi : \mathbb{R}_0^+ \longrightarrow \mathbb{R}_0^+$ satisfying*

$$w_F(y) \leq \beta \Phi(y^{-\alpha})^{-\gamma}, \tag{10}$$

$$\kappa_i < \gamma\left(\frac{i}{r} + 2(p-i) - d\right), \quad i = 1, \ldots, p, \tag{11}$$

$$\Phi(2z) \leq \lambda \Phi(z), \quad \forall z \geq z_0, \tag{12}$$

such that for all $f \in \mathcal{H}_\alpha$ with $\|f\|_\infty \leq 1$, the moment bound (2) holds with $\Phi_i = \Phi^{\kappa_i}$.

Then there is a centered Gaussian process $(W(t))_{t \in [-\infty,\infty]^d}$ with almost surely continuous sample paths such that $U_n \xrightarrow{\mathcal{D}} W$, in the space $\mathbb{D}([-\infty,\infty]^d)$.

Proof. Using Lemma 1, it is sufficient to show condition (9) of Theorem 1 for the function Ψ defined in (7) and for suitable $\gamma_1, \ldots, \gamma_p$ satisfying (8).

By (10), for $z > 0$ we have

$$w_F^\leftarrow\left(\frac{1}{z}\right) \geq \inf\left\{x > 0 : \beta(\Phi(x^{-\alpha}))^{-\gamma} \geq \frac{1}{z}\right\} = \inf\left\{x > 0 : \Phi(x^{-\alpha}) \leq (\beta z)^{\frac{1}{\gamma}}\right\}$$

and therefore, since Φ is invertible, we obtain

$$w_F^\leftarrow\left(\frac{1}{z}\right)^{-\alpha} \geq \inf\left\{x > 0 : x \geq \left(\Phi^{-1}((\beta z)^{\frac{1}{\gamma}})\right)^{-\frac{1}{\alpha}}\right\}^{-\alpha} = \Phi^{-1}\left((\beta z)^{\frac{1}{\gamma}}\right).$$

Using (12), this yields that there is a finite constant $C' \in \mathbb{R}$ such that

$$\begin{aligned}
\Phi_i(2\Psi(z)) &= \Phi_i\left(2d\left(w_F^\leftarrow\left(\frac{1}{z}\right)\right)^{-\alpha} + 2\right) \\
&\leq \left(\Phi(2d\Phi^{-1}((\beta z)^{\frac{1}{\gamma}}) + 2)\right)^{\kappa_i} \\
&\leq C'\left(\Phi(\Phi^{-1}((\beta z)^{\frac{1}{\gamma}}))\right)^{\kappa_i} \\
&= C'\beta^{\frac{\kappa_i}{\gamma}} \cdot z^{\frac{\kappa_i}{\gamma}}
\end{aligned}$$

for every $i \in \{1, \ldots, p\}$ and all sufficiently large z. Here, by (11), each exponent $\gamma_i = \kappa_i/\gamma$ is positive and strictly smaller than $i/r + 2(p-i) - d$. Thus conditions (9) and (8) of Theorem 1 are satisfied. $\qquad\square$

Remark 2. We will now show that by taking $\Phi(x) = \log(x+1)$ and special choices of the constants, the main results of [7] and [8] can be obtained as corollaries of Theorem 3. Note that this Φ-function arises in (2) when the process is multiple mixing with exponential decay (see [7]). For the above choice of Φ, (10) can be simplified to

$$w_F(y) \leq \beta |\log(y)|^{-\gamma}$$

for some $\beta > 0$.

1) If we choose $\kappa_i = 2p - i$, we get

$$\max_{i=1,\ldots,p} \frac{2p-i}{\frac{i}{r} + 2(p-i) - d} = \frac{rp}{p-rd},$$

and thus condition (11) in Theorem 3 is equivalent to $\gamma > \frac{rp}{p-rd}$. In this way, we obtain Theorem 3 in [7, p.1081].

2) If we choose $\alpha, d, r = 1$ and $p = 2$, condition (11) in Theorem 3 reduces to $\gamma > \max\{\frac{\kappa_1}{2}, \kappa_2\}$. Thus we obtain Theorem 1 in [8, p.3702].

This shows that Theorem 3 is a far-reaching generalization of the result obtained in [7] and [8].

We finally obtain a version of Theorem 2 for the special choice $\mathcal{G} = \mathcal{H}_\alpha$.

Corollary 1. *If conditions* (i) *and* (ii) *of Theorem 2 hold with* $\mathcal{G} = \mathcal{H}_\alpha$ *and if F is ϑ-Hölder for some ϑ such that*

$$\frac{\vartheta}{\alpha} > \frac{rp}{p-rd}, \tag{13}$$

then the empirical CLT holds.

Proof. We will apply Theorem 3 and thus have to verify assumption (ii). Choosing $\kappa_i = i$ and $\Phi(x) = x$, we obtain the moment bound (2) by Proposition 1, while (12) is obviously satisfied. Choosing moreover $\gamma = \frac{\vartheta}{\alpha}$, (11) follows from (13). Finally (10) is a consequence of the Hölder assumption on F. $\square$

3.2 Causal functions of independent and identically distributed processes

One example of processes that feature the multiple mixing property (3) and that can be treated by our methods is the class of causal functions of i.i.d. processes, which are defined as follows:

Definition 3 (Causal function). Let $(\xi_j)_{j\in\mathbb{Z}}$ be an independent identically distributed process with values in a Banach space $\mathcal{X}$. We call $(X_i)_{i\in\mathbb{N}}$ a causal function of $(\xi_j)_{j\in\mathbb{Z}}$ if there is a measurable function $G : \mathcal{X}^\mathbb{N} \to \mathbb{R}^d$ such that each X_i is of the form

$$X_i := G((\xi_{i-j})_{j\in\mathbb{N}}).$$

Let us now introduce a measure of the dependence structure of a causal function of an i.i.d. process $(\xi_j)_{j\in\mathbb{Z}}$. Set

$$\dot{X}_i := G(\xi_i, \xi_{i-1}, \ldots, \xi_1, \xi'_0, \xi'_{-1}, \ldots),$$

where $(\xi'_j)_{j\in\mathbb{Z}}$ is an independent copy of $(\xi_j)_{j\in\mathbb{Z}}$, i.e. $(\xi_j)_{j\in\mathbb{Z}}$ and $(\xi'_j)_{j\in\mathbb{Z}}$ are identically distributed and both processes are independent from each other. We can now define for $i \in \mathbb{N}^*$ and $s \geq 1$,

$$\delta_{i,s} = \|X_i - \dot{X}_i\|_s := \mathrm{E}\left(\|X_i - \dot{X}_i\|_{\mathbb{R}^d}^s\right)^{\frac{1}{s}}. \tag{14}$$

This physical dependence measure was introduced by Wu [18].

Proposition 2. *Let $(X_i)_{i \in \mathbb{N}}$ be an $\mathbb{R}^d$-valued causal function of an i.i.d. process, $\alpha \in (0, 1]$ and $r \in [1, \infty)$, $s \in (1, \infty]$ with $\frac{1}{r} + \frac{1}{s} = 1$. Then*

$$(X_i)_{i \in \mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{H}_\alpha),$$

where $\Theta(i) = (\delta_{i,s})^\alpha$. As a consequence of Proposition 1, if

$$\sum_{i=1}^{\infty} i^{2p-2}(\delta_{i,s})^\alpha < \infty \tag{15}$$

for some $p > rd$, then the moment bound (2) holds for all $f \in \mathcal{H}_\alpha$ such that $\|f\|_\infty \le 1$, with p, r as above and $\Phi_i(x) = x^i$.

The second crucial point in the assumptions of Theorem 1 is the CLT (1). The following proposition gives a criterion for (1) in the situation of a causal function of an i.i.d. process.

Proposition 3. *If $(X_i)_{i \in \mathbb{N}}$ is an $\mathbb{R}^d$-valued causal function of an i.i.d. process and satisfies*

$$\sum_{i=1}^{\infty} (\delta_{i,s})^\alpha < \infty, \tag{16}$$

for some $\alpha \in (0, 1]$, $s \in [1, \infty]$, then the CLT (1) under $\mathcal{H}_\alpha$ holds with $\sigma_f^2 = \mathrm{E}(f(X_0)^2) + 2 \sum_{i=1}^{\infty} \mathrm{E}\big(f(X_0)f(X_i)\big)$.

The proofs of the two preceding propositions are given in Section 5. As a direct application of previous results, we obtain the following one:

Theorem 4. *Let $(X_i)_{i \in \mathbb{N}}$ be an $\mathbb{R}^d$-valued causal function of an i.i.d. sequence. Assume that:*

(i) The distribution function F of X_0 is ϑ-Hölder for some $\vartheta \in (0, 1]$.

(ii) There are some $r \in [1, \infty)$, $s \in (1, \infty]$ satisfying $\frac{1}{r} + \frac{1}{s} = 1$, an integer $p > rd$ and a positive constant $\alpha \in (0, 1]$ satisfying (13) and (15).

Then there is a centered Gaussian process $(W(t))_{t \in [-\infty, \infty]^d}$ with almost surely continuous sample paths such that $U_n \xrightarrow{\mathcal{D}} W$, in the space $\mathbb{D}([-\infty, \infty]^d)$.

Proof. We will apply Corollary 1. By Proposition 3 the CLT (1) holds under $\mathcal{H}_\alpha$. Proposition 2 shows that $(X_i)_{i \in \mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{H}_\alpha)$, where $\Theta(i) = (\delta_{i,s})^\alpha$ and thus $\sum_{i=0}^{\infty} i^{2p-2}\Theta(i) < \infty$. $\square$

Example 2 (Time delay vectors). Let $(X_i)_{i \in \mathbb{N}}$ be a real-valued causal function of an i.i.d. process. We define the time delay vector process $(Y_i)_{i \in \mathbb{N}}$ of dimension $d \ge 1$, by

$$Y_i = (X_i, \dots, X_{i+d-1}), \quad i \in \mathbb{N}.$$

If the scalar process $(X_i)_{i \in \mathbb{N}}$ satisfies (i) and (ii) of Theorem 4, then the Empirical Central Limit Theorem holds for the process $(Y_i)_{i \in \mathbb{N}}$.

Proof. Assume that $(X_i)_{i \in \mathbb{N}}$ satisfies (i) and (ii) of Theorem 4 and let us check that the process $(Y_i)_{i \in \mathbb{N}}$ also satisfies these assumptions. Denote by F_X the distribution function of X_0 and by F_Y the multidimensional distribution function of Y_0. The marginals of F_Y are all F_X and then $w_{F_Y} \le d w_{F_X}$. Thus F_Y is ϑ-Hölder. Denote by $\delta_{i,s}(X)$ and $\delta_{i,s}(Y)$ the coefficients introduced in (14) relative respectively to $(X_i)_{i \in \mathbb{N}}$ and $(Y_i)_{i \in \mathbb{N}}$. We can see that there exists a constant $C > 0$ such that for all $i \in \mathbb{N}$,

$$\delta_{i,s}(Y) \le C(\delta_{i,s}(X) + \dots + \delta_{i+d-1,s}(X)).$$

Thus, we infer that $(Y_i)_{i \in \mathbb{N}}$ satisfies (15) with the same constant α as for $(X_i)_{i \in \mathbb{N}}$. $\square$

Example 3 (Linear processes)**.** Let $(X_i)_{i \in \mathbb{N}}$ be a causal linear function given by

$$X_i := \sum_{j \in \mathbb{N}} a_j \xi_{i-j},$$

where $(\xi_j)_{j \in \mathbb{Z}}$ is an i.i.d. $\mathcal{X}$-valued process and $(a_j)_{j \in \mathbb{N}}$ is a family of linear operators from $\mathcal{X}$ to $\mathbb{R}^d$. We denote the norm of such operators by

$$|a| = \sup\{\|a(x)\| : x \in \mathcal{X}, |x|_{\mathcal{X}} \leq 1\}.$$

If $\|\xi_0\|_s < \infty$ for some $s > 1$, if the distribution function F of X_0 is ϑ-Hölder and if

$$\sum_{j=i}^{\infty} |a_j| = \mathcal{O}(i^{-b}) \ \text{ with } b > \min_{p \in \mathbb{N}, p > rd} \frac{r}{\vartheta} \frac{(2p-1)p}{p-rd}$$

for $r = \frac{s}{s-1}$, then the Empirical Central Limit Theorem holds.

Proof. Let $(\xi'_j)_{j \in \mathbb{Z}}$ be an independent copy of $(\xi_j)_{j \in \mathbb{Z}}$ and p an integer which realizes the minimum in the condition on b. By assumption, there is an $\varepsilon > 0$ such that $b > (1+\varepsilon)\frac{(2p-1)rp}{\vartheta(p-rd)}$. We can choose $\alpha = \vartheta(1+\varepsilon)^{-1}\frac{p-rd}{rp}$, ensuring that (13) is satisfied. Since

$$\delta_{i,s} = \left\| \sum_{j=i}^{\infty} a_j(\xi_{i-j} - \xi'_{i-j}) \right\|_s \leq \|\xi_0 - \xi'_0\|_s \sum_{j=i}^{\infty} |a_j|,$$

we have

$$i^{2p-2}(\delta_{i,s})^{\alpha} \leq (2\|\xi_0\|_s)^{\alpha} i^{2p-2} \Big(\sum_{j=i}^{\infty} |a_j| \Big)^{\alpha} = \mathcal{O}(i^{2p-2-\alpha b}),$$

where $2p - 2 - \alpha b < 2p - 2 - (2p - 1) = -1$, thus (15) holds and Theorem 4 applies. $\qquad \square$

The example of causal linear processes has already been studied by several authors. Dedecker and Prieur [6] could allow lower rates of convergence for $\sum_{j=i}^{\infty} |a_j|$ but required that X_0 has a bounded density. Wu [19] also assumes that the underlying i.i.d. process has a density. In the present paper, no assumption is made on the distribution of the ξ_i except moments and the distribution of X_0 does not need to be absolutely continuous. In the case where ξ_i are $\mathbb{R}$-valued and $d = 1$, very weak conditions can be found in [5].

4 Proof of Theorem 1

To prove the convergence of the empirical process we will use the following result obtained in [8, Theorem 2]. Note that this result is a modification of a classical theorem on weak convergence; see [2, Theorem 4.2].

Proposition 4. *Let S be a separable and complete space with metric ρ. If $U_n, U_n^{(m)}$ and $U^{(m)}$, $n, m \in \mathbb{N}^*$, are S-valued random variables such that*

$$U_n^{(m)} \xrightarrow{\mathcal{D}} U^{(m)} \text{ as } n \to \infty, \text{ for all } m \in \mathbb{N}^* \text{ and} \tag{17}$$

$$\lim_{m \to \infty} \limsup_{n \to \infty} \mathrm{P}(\rho(U_n, U_n^{(m)}) > \varepsilon) = 0 \text{ for all } \varepsilon > 0, \tag{18}$$

then there is an S-valued random variable U such that $U_n \xrightarrow{\mathcal{D}} U$ as $n \to \infty$ and $U^{(m)} \xrightarrow{\mathcal{D}} U$ as $m \to \infty$.

Remark 3. We will apply Proposition 4 to the situation where S is chosen as $\mathbb{D}([-\infty, \infty]^d)$ (equipped with the Skorokhod metric d_0, c.f. p.2) and U_n denotes the empirical process. If one wants to prove the convergence in distribution of U_n Proposition 4 assures us, that it is sufficient to find a process $U_n^{(m)}$ which approximates U_n as $m \to \infty$ in the sense of (18) and to show that this process is convergent in distribution for each m as $n \to \infty$.

Following the techniques presented in [7, p.1078 ff], we begin by introducing a partition for $[-\infty, \infty]^d$. Let F_i be the i-th marginal distribution of F, $0 = r_0^{(m)} < r_1^{(m)} < \ldots < r_m^{(m)} = 1$ a partition of $[0, 1]$ and set for $i \in \{1, \ldots, d\}$, $j_i \in \{0, \ldots, m\}$,

$$t_{i,j_i}^{(m)} := F_i^{\rightarrow}(r_{j_i}^{(m)}),$$

where $F_i^{\rightarrow}(y) := \sup\{x \in [-\infty, \infty] : F_i(x) \leq y\}$.

Note that $F_i^{\rightarrow}$ are injective since F_i are continuous. For convenience we also define $t_{i,m+1}^{(m)} := t_{i,m}^{(m)}$. For $j \in \{0, \ldots, m+1\}^d$ set

$$t_j^{(m)} := (t_{1,j_1}^{(m)}, \ldots, t_{d,j_d}^{(m)}) = (F_1^{\rightarrow}(r_{j_1}^{(m)}), \ldots, F_d^{\rightarrow}(r_{j_d}^{(m)})).$$

To keep notation short, denote $(x, \ldots, x) \in [-\infty, \infty]^d$ by $\overline{x}$.

We can construct a $\mathcal{G}$-approximation of the indicator function $\mathbf{1}_{[-\overline{\infty}, t_{j-\overline{1}}^{(m)}]}$ by setting for $j \in \{1, \ldots, m\}^d$

$$\varphi_j^{(m)} := \begin{cases} \varphi_{(t_{j-\overline{2}}^{(m)}, t_{j-\overline{1}}^{(m)})}, & \text{if } j \geq \overline{2}, \\ 0, & \text{if } j \not\geq \overline{2}, \end{cases} \tag{19}$$

where $\varphi_{(t_{j-\overline{2}}^{(m)}, t_{j-\overline{1}}^{(m)})} \in \mathcal{G}$ satisfies (5) and (6). Observe that $t_{j-\overline{2}}^{(m)} < t_{j-\overline{1}}^{(m)}$, since all $F_i^{\rightarrow}$ are injective.

To approximate the empirical distribution function we introduce

$$F_n^{(m)}(t) := \sum_{j \in \{1, \ldots, m\}^d} \left(\frac{1}{n} \sum_{i=1}^{n} \varphi_j^{(m)}(X_i) \right) \mathbf{1}_{[t_{j-\overline{1}}^{(m)}, t_j^{(m)})}(t).$$

Note that for t in any fixed rectangle $[t_{j-\overline{1}}^{(m)}, t_j^{(m)})$ we have the simple form

$$F_n^{(m)}(t) = \frac{1}{n} \sum_{i=1}^{n} \varphi_j^{(m)}(X_i).$$

By the definition of the $\varphi_j^{(m)}$ it is easy to see that therefore

$$F_n(t_{j-\overline{2}}^{(m)}) \leq F_n^{(m)}(t) \leq F_n(t_{j-\overline{1}}^{(m)}) \quad \forall t \in [t_{j-\overline{1}}^{(m)}, t_j^{(m)}).$$

Thus it is natural to approximate (as $m \to \infty$) U_n by

$$U_n^{(m)} := \left(\sqrt{n} \left(F_n^{(m)}(t) - F^{(m)}(t) \right) \right)_{t \in [-\infty, \infty]^d},$$

where

$$F^{(m)}(t) := \mathrm{E}\left(F_n^{(m)}(t) \right) = \sum_{j \in \{1, \ldots, m\}^d} \mathrm{E}\left(\varphi_j^{(m)}(X_0) \right) \mathbf{1}_{[t_{j-\overline{1}}^{(m)}, t_j^{(m)})}(t).$$

Remark 4. Notice that at this point the $\varphi_j^{(m)}$ and thus $U_n^{(m)}$ depend heavily on the chosen partition $r_0, \ldots, r_m$ on $[0, 1]$. Therefore the notation with the superscript m may be misleading at first glance, but since whenever the choice of the partition matters we will only use equidistant partitions of $[0, 1]$, the partitions will in all relevant situations be uniquely defined by m.

The central idea to prove Theorem 1 is to use Proposition 4. Thus we need to check (17) and (18) for $S = \mathbb{D}([-\infty, \infty]^d)$. This is done in the next two lemmas.

Lemma 2. *For every partition $0 = r_0^{(m)} < \ldots < r_m^{(m)} = 1$ of $[0, 1]$, $U_n^{(m)}$ converges weakly to some centered Gaussian process $W^{(m)} \in \mathbb{D}([-\infty, \infty]^d)$ whose sample paths are constant on each of the rectangles $[t_{j-1}^{(m)}, t_j^{(m)})$, $j \in \{1, \ldots, m\}^d$.*

Proof. Since all the $U_n^{(m)}$ are constant on each of the rectangles $[t_{j-1}^{(m)}, t_j^{(m)})$, it suffices to show weak convergence of the sequence of vectors

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\varphi_j^{(m)}(X_i) - \mathrm{E}\left(\varphi_j^{(m)}(X_i) \right) \right) \right)_{j \in \{1, \ldots, m\}^d},$$

which is a consequence of the CLT (1) and the Cramér-Wold device. $\square$

Lemma 3. *Let $0 = r_0^{(m)} < r_1^{(m)} < \ldots < r_m^{(m)} = 1$ be the partition of $[0, 1]$ defined by $r_k^{(m)} = \frac{k}{m}$. Then for every $\varepsilon, \eta > 0$ there is a $m_0 \in \mathbb{N}^*$ such that for all $m \geq m_0$*

$$\limsup_{n \to \infty} \mathrm{P}\left(\sup_{t \in [-\infty, \infty]^d} |U_n(t) - U_n^{(m)}(t)| > \varepsilon \right) \leq \eta.$$

Proof. Let us consider $\varepsilon, \eta > 0$ fixed for the rest of this proof. Consider the partition $0 = r_0^{(m)} < \ldots < r_m^{(m)} = 1$ of $[0, 1]$ defined in the statement of the lemma and set $h = \frac{1}{m}$. For each $k \in \mathbb{N}^*$, consider the refined partition

$$r_{q-1}^{(m)} = s_{q,0}^{(k)} < s_{q,1}^{(k)} < \ldots < s_{q,2^k}^{(k)} = r_q^{(m)}$$

of $[r_{q-1}^{(m)}, r_q^{(m)}]$, where

$$s_{q,\ell}^{(k)} := r_{q-1}^{(m)} + \ell \cdot \frac{h}{2^k},$$

$\ell \in \{0, \ldots, 2^k\}$ and $q \in \{0, \ldots, m\}$. Setting for $i \in \{1, \ldots, d\}$, $j_i \in \{1, \ldots, m\}$, $l_i \in \{0, \ldots, 2^k\}$

$$s_{i,j_i,l_i}^{(k)} = F_i^{\rightarrow}(s_{j_i,l_i}^{(k)}),$$

we obtain partitions

$$t_{i,j_i-1}^{(m)} = s_{i,j_i,0}^{(k)} < s_{i,j_i,1}^{(k)} < \ldots < s_{i,j_i,2^k}^{(k)} = t_{i,j_i}^{(m)}$$

of $[t_{i,j_i-1}^{(m)}, t_{i,j_i}^{(m)}]$. To simplify the notation in the following calculations we set

$$s_{i,j_i,-1}^{(k)} := s_{i,j_i-1,2^k-1}^{(k)}, \quad \text{for } j_i > 1, \quad \text{and} \quad s_{i,j_i,2^k+1}^{(k)} := s_{i,j_i+1,1}^{(k)}, \quad \text{for } j_i < m.$$

Let us now focus on a fixed rectangle $[t_{j-1}^{(m)}, t_j^{(m)})$ for some $j = (j_1, \ldots, j_d) \in \{1, \ldots, m\}^d$. Our aim is to construct a chain to link the point $t_{j-1}^{(m)}$ to some arbitrary point $t \in [t_{j-1}^{(m)}, t_j^{(m)})$. Therefore we set

$$l_{i,j_i}(k, t) = \max\{\ell \in \{0, \ldots, 2^k\} : s_{i,j_i,\ell}^{(k)} \leq t_i\} \in \{0, \ldots, 2^k - 1\}.$$

Since we consider j to be fixed, we may drop the index j in order to simplify further notation. More precisely, we set

$$s_l^{(k)} := (s_{1,j_1,l_1}^{(k)}, \ldots, s_{d,j_d,l_d}^{(k)}), \quad \text{and} \quad l(k,t) := (l_{1,j_1}(k,t), \ldots, l_{d,j_d}(k,t)).$$

In this way for any $k \in \mathbb{N}^*$, we obtain an $([-\infty, \infty]^d$- valued) chain

$$t_{j-\overline{1}}^{(m)} = s_{l(0,t)}^{(0)} \le s_{l(1,t)}^{(1)} \le \ldots \le s_{l(k,t)}^{(k)} \le t \le s_{l(k,t)+\overline{1}}^{(k)}.$$

Now set

$$\psi_{\overline{0}}^{(0)} := \varphi_{(t_{j-\overline{1}}^{(m)}, t_j^{(m)})}$$

and choose for every $k \in \mathbb{N}^*$ and $l \in \{0, \ldots, 2^k + 1\}^d$ a function $\psi_l^{(k)} \in \mathcal{G}$ such that[4]

$$\psi_l^{(k)} = \begin{cases} 0, & \text{if } \exists i \in \{1, \ldots, d\} : j_i = 1 \text{ and } l_i = 0 \\ 1, & \text{if } (\exists i \in \{1, \ldots, d\} : j_i = m \text{ and } l_i = 2^k + 1) \\ & \text{and } (\nexists i \in \{1, \ldots, d\} : j_i = 1 \text{ and } l_i = 0) \\ \varphi_{(s_{l-\overline{1}}^{(k)}, s_l^{(k)})} & \text{else,} \end{cases} \tag{20}$$

where $\varphi_{(s_{l-\overline{1}}^{(k)}, s_l^{(k)})}$ satisfies (5) and (6).

By this definition we have for every $t \in [-\infty, \infty]^d$ and $l \in \{0, \ldots, 2^k\}^d$ the following inequalities:

$$\mathbf{1}_{[-\overline{\infty}, s_{l-\overline{1}}]} \le \psi_l^{(k)} \le \mathbf{1}_{[-\overline{\infty}, s_l]} \tag{21}$$

$$\varphi_j^{(m)} \le \psi_{l(1,t)}^{(1)} \le \ldots \le \psi_{l(k,t)}^{(k)} \le \mathbf{1}_{[-\overline{\infty}, t]} \le \psi_{l(k,t)+\overline{2}}^{(k)}. \tag{22}$$

Using inequality (22) we obtain for $t \in [t_{j-\overline{1}}^{(m)}, t_j^{(m)})$ and $K \in \mathbb{N}^*$, the telescopic-sum representation

$$\frac{1}{n} \sum_{i=1}^n \left(\mathbf{1}_{[-\overline{\infty}, t]}(X_i) - F_n^{(m)}(t) \right)$$

$$= \sum_{k=1}^K \frac{1}{n} \sum_{i=1}^n \left(\psi_{l(k,t)}^{(k)}(X_i) - \psi_{l(k-1,t)}^{(k-1)}(X_i) \right) + \frac{1}{n} \sum_{i=1}^n \left(\mathbf{1}_{[-\overline{\infty}, t]}(X_i) - \psi_{l(K,t)}^{(K)}(X_i) \right). \tag{23}$$

Let us now consider

$$U_n(t) - U_n^{(m)}(t) = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{[-\overline{\infty}, t]}(X_i) - F(t) \right) - \sqrt{n} \left(F_n^{(m)}(t) - F^{(m)}(t) \right).$$

Equation (23) yields

$$U_n(t) - U_n^{(m)}(t)$$

$$= \sum_{k=1}^K \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\psi_{l(k,t)}^{(k)}(X_i) - \mathrm{E}\psi_{l(k,t)}^{(k)}(X_0) \right) - \left(\psi_{l(k-1,t)}^{(k-1)}(X_i) - \mathrm{E}\psi_{l(k-1,t)}^{(k-1)}(X_0) \right)$$

$$+ \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\mathbf{1}_{[-\overline{\infty}, t]}(X_i) - F(t) \right) - \left(\psi_{l(K,t)}^{(K)}(X_i) - \mathrm{E}\psi_{l(K,t)}^{(K)}(X_0) \right). \tag{24}$$

---

[4]the reference to the indices j and m is omitted, since these are considered to be fixed.

Applying the inequalities in (22), we gain the following upper bounds for the last sum on the right-hand side of the above inequality:

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\left(\mathbf{1}_{[-\infty,t]}(X_i) - F(t) \right) - \left(\psi_{l(K,t)}^{(K)}(X_i) - \mathrm{E}\psi_{l(K,t)}^{(K)}(X_0) \right) \right)$$
$$\geq -\sqrt{n}\left(\mathrm{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0) - \mathrm{E}\psi_{l(K,t)}^{(K)}(X_0) \right) \tag{25}$$

and

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\left(\mathbf{1}_{[-\infty,t]}(X_i) - F(t) \right) - \left(\psi_{l(K,t)}^{(K)}(X_i) - \mathrm{E}\psi_{l(K,t)}^{(K)}(X_0) \right) \right)$$
$$\leq \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\left(\psi_{l(K,t)+\overline{2}}^{(K)}(X_i) - \mathrm{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0) \right) - \left(\psi_{l(K,t)}^{(K)}(X_i) - \mathrm{E}\psi_{l(K,t)}^{(K)}(X_0) \right) \right)$$
$$+ \sqrt{n}\left(\mathrm{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0) - \mathrm{E}\psi_{l(K,t)}^{(K)}(X_0) \right). \tag{26}$$

For convenience, let $s_{i,m,2^k+1}^{(k)} := s_{i,m,2^k}^{(k)}$. By equation (21) and the continuity[5] of F, we obtain

$$\sqrt{n}\left| \mathrm{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0) - \mathrm{E}\psi_{l(K,t)}^{(K)}(X_0) \right|$$
$$\leq \sqrt{n}\left| \mathrm{E}\mathbf{1}_{[-\infty,s_{l(K,t)+\overline{2}})}(X_0) - \mathrm{E}\mathbf{1}_{[-\infty,s_{l(K,t)-\overline{1}})}(X_0) \right|$$
$$\leq \sqrt{n}\left(d \max_{i=1,\ldots,d} \left\{ F_i(s_{i,j_i,l_{i,j_i}(K,t)+\overline{2}}) - F_i(s_{i,j_i,l_{i,j_i}(K,t)-\overline{1}}) \right\} \right) = \frac{3d\sqrt{n}h}{2^K},$$

and thus, if we choose

$$K = K_n := \left\lfloor \log_2\left(\frac{2^4 d}{\varepsilon} \sqrt{n}h \right) \right\rfloor, \tag{27}$$

we obtain

$$\left| \mathrm{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0) - \mathrm{E}\psi_{l(K,t)}^{(K)}(X_0) \right| < \frac{\varepsilon}{2}. \tag{28}$$

In summary, using (25), (26) and (28) in equation (24) yields, for all $n \in \mathbb{N}^*$,

$$\left| U_n(t) - U_n^{(m)}(t) \right| < \left| \sum_{k=1}^{K_n} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \left(\psi_{l(k,t)}^{(k)}(X_i) - \mathrm{E}\psi_{l(k,t)}^{(k)}(X_0) \right) \right. \right.$$
$$\left. - \left(\psi_{l(k-1,t)}^{(k-1)}(X_i) - \mathrm{E}\psi_{l(k-1,t)}^{(k-1)}(X_0) \right) \right\} \Bigg|$$
$$+ \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \left(\psi_{l(K_n,t)+\overline{2}}^{(K_n)}(X_i) - \mathrm{E}\psi_{l(K_n,t)+\overline{2}}^{(K_n)}(X_0) \right) \right. \right.$$
$$\left. \left. - \left(\psi_{l(K_n,t)}^{(K_n)}(X_i) - \mathrm{E}\psi_{l(K_n,t)}^{(K_n)}(X_0) \right) \right\} \right| + \frac{\varepsilon}{2}. \tag{29}$$

Consider the maximum of the terms in (29) over all $t \in [t_{j-\overline{1}}^{(m)}, t_j^{(m)})$. By the definition of the $l(k,t)$ we have

$$\left\lfloor \frac{l(k,t)}{2} \right\rfloor := \left(\left\lfloor \frac{l_{1,j_1}(k,t)}{2} \right\rfloor, \ldots, \left\lfloor \frac{l_{d,j_d}(k,t)}{2} \right\rfloor \right) = l(k-1,t).$$

---

[5]note that for continuous F, we have $F \circ F^{\rightarrow}(x) = x$ for all $x \in [0,1]$.

We therefore obtain

$$
\sup_{t \in [t_{j-1}^{(m)}, t_j^{(m)})} |U_n(t) - U_n^{(m)}(t)| < \sum_{k=1}^{K_n} \frac{1}{\sqrt{n}} \max_{l \in \{0,\ldots,2^k-1\}^d} \left| \sum_{i=1}^n \left\{ \left(\psi_l^{(k)}(X_i) - \mathrm{E}\psi_l^{(k)}(X_0) \right) \right. \right.
$$

$$
\left. \left. - \left(\psi_{\lfloor l/2 \rfloor}^{(k-1)}(X_i) - \mathrm{E}\psi_{\lfloor l/2 \rfloor}^{(k-1)}(X_0) \right) \right\} \right|
$$

$$
+ \frac{1}{\sqrt{n}} \max_{l \in \{0,\ldots,2^{K_n}-1\}^d} \left| \sum_{i=1}^n \left\{ \left(\psi_{l+\frac{1}{2}}^{(K_n)}(X_i) - \mathrm{E}\psi_{l+\frac{1}{2}}^{(K_n)}(X_0) \right) \right. \right.
$$

$$
\left. \left. - \left(\psi_l^{(K_n)}(X_i) - \mathrm{E}\psi_l^{(K_n)}(X_0) \right) \right\} \right|
$$

$$
+ \frac{\varepsilon}{2}.
$$

Choose $\varepsilon_k = \frac{\varepsilon}{4k(k+1)}$ and note that $\sum_{k=1}^\infty \varepsilon_k = \varepsilon/4$. An application of Markov's inequality for the $2p$-th moments combined with condition (2) implies

$$
\mathrm{P}\left(\sup_{t \in [t_{j-1}^{(m)}, t_j^{(m)})} |U_n(t) - U_n^{(m)}(t)| \geq \varepsilon \right)
$$

$$
\leq \sum_{k=1}^{K_n} \sum_{l \in \{0,\ldots,2^k-1\}^d} \mathrm{P}\left(\frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \left\{ \left(\psi_l^{(k)}(X_i) - \mathrm{E}\psi_l^{(k)}(X_0) \right) \right. \right. \right.
$$

$$
\left. \left. \left. - \left(\psi_{\lfloor l/2 \rfloor}^{(k-1)}(X_i) - \mathrm{E}\psi_{\lfloor l/2 \rfloor}^{(k-1)}(X_0) \right) \right\} \right| > \varepsilon_k \right)
$$

$$
+ \sum_{l \in \{0,\ldots,2^{K_n}-1\}^d} \mathrm{P}\left(\frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \left\{ \left(\psi_{l+\frac{1}{2}}^{(K_n)}(X_i) - \mathrm{E}\psi_{l+\frac{1}{2}}^{(K_n)}(X_0) \right) \right. \right. \right.
$$

$$
\left. \left. \left. - \left(\psi_l^{(K_n)}(X_i) - \mathrm{E}\psi_l^{(K_n)}(X_0) \right) \right\} \right| > \frac{\varepsilon}{4} \right)
$$

$$
\leq 2C\left\{ \sum_{k=1}^{K_n} \sum_{l \in \{0,\ldots,2^k-1\}^d} \frac{1}{\varepsilon_k^{2p} n^p} \right.
$$

$$
\cdot \sum_{i=1}^p n^i \, \|\psi_l^{(k)}(X_0) - \psi_{\lfloor l/2 \rfloor}^{(k-1)}(X_0)\|_r^i \, \Phi_i\left(2\|\psi_l^{(k)} - \psi_{\lfloor l/2 \rfloor}^{(k-1)}\|_{\mathcal{G}} \right)
$$

$$
+ \sum_{l \in \{0,\ldots,2^{K_n}-1\}^d} \frac{4^{2p}}{\varepsilon^{2p} n^p}
$$

$$
\left. \cdot \sum_{i=1}^p n^i \, \|\psi_{l+\frac{1}{2}}^{(K_n)}(X_0) - \psi_l^{(K_n)}(X_0)\|_r^i \, \Phi_i\left(2\|\psi_{l+\frac{1}{2}}^{(K_n)} - \psi_l^{(K_n)}\|_{\mathcal{G}} \right) \right\}. \tag{30}
$$

The critical part in these terms is the argument in the functions Φ_i. We therefore collect the necessary auxiliary calculations in the following lemma:

15

Lemma 4. *For all $l \in \{0, \ldots, 2^k - 1\}^d$, $k, n, r \in \mathbb{N}^*$*

$$\|\psi_l^{(k)}(X_0) - \psi_{\lfloor l/2 \rfloor}^{(k-1)}(X_0)\|_r \leq \left(\frac{3dh}{2^k}\right)^{\frac{1}{r}},$$

$$\|\psi_{l+\overline{2}}^{(K_n)}(X_0) - \psi_l^{(K_n)}(X_0)\|_r \leq \left(\frac{3dh}{2^{K_n}}\right)^{\frac{1}{r}},$$

$$\|\psi_l^{(k)}\|_{\mathcal{G}} \leq \max\left\{\Psi\left(\frac{2^k}{h}\right), \|1\|_{\mathcal{G}}\right\}$$

Proof. By (21) and the continuity of the F_i,

$$\|\psi_l^{(k)}(X_0) - \psi_{\lfloor l/2 \rfloor}^{(k-1)}(X_0)\|_r \leq \|\mathbf{1}_{[-\overline{\infty}, s_l^{(k)}]}(X_0) - \mathbf{1}_{[-\overline{\infty}, s_{\lfloor l/2 \rfloor - \overline{1}}^{(k-1)}]}(X_0)\|_r$$

$$\leq \left(d \max_{i=1,\ldots,d}\left(F_i(s_{i,j_i,l_i}^{(k)}) - F_i(s_{i,j_i,l_i-3}^{(k)})\right)\right)^{\frac{1}{r}}$$

$$\leq \left(\frac{3dh}{2^k}\right)^{\frac{1}{r}}.$$

The second inequality can be proven in a similar way.

In the first two cases of the definition (20), $\psi_l^{(k)}$ is a constant function taking either the value zero or one for each argument. In this cases the last inequality of the lemma is trivially satisfied by the conditions on $\|\cdot\|_{\mathcal{G}}$. Else $\psi_l^{(k)}$ has a representation $\varphi_{(s_{l-\overline{1}}^{(k)}, s_l^{(k)})}$, where

$$s_l^{(k)} = \left(s_{1,j_1,l_1}^{(k)}, \ldots, s_{d,j_d,l_d}^{(k)}\right) = \left(F_1^{\rightarrow}(s_{j_1,l_1}^{(k)}), \ldots, F_d^{\rightarrow}(s_{j_d,l_d}^{(k)})\right),$$

$$s_{l-\overline{1}}^{(k)} = \left(s_{1,j_1,l_1-1}^{(k)}, \ldots, s_{d,j_d,l_d-1}^{(k)}\right)$$

$$= \left(F_1^{\rightarrow}(s_{j_1,l_1}^{(k)} - h2^{-k}), \ldots, F_d^{\rightarrow}(s_{j_d,l_d}^{(k)} - h2^{-k})\right)$$

and hence, for every $i \in \{1, \ldots, d\}$,

$$s_{i,j_i,l_i}^{(k)} - s_{i,j_i,l_i-1}^{(k)} \in \{\delta > 0 : \exists t \in \mathbb{R}, \ |F_i(t) - F_i(t - \delta)| \geq h2^{-k}\}$$

$$\subset \{\delta > 0 : w_{F_i}(\delta) \geq h2^{-k}\} \tag{31}$$

To see this, set $\delta = F_i^{\rightarrow}(s_{l_i}^{(k)}) - F_i^{\rightarrow}(s_{l_i-1}^{(k)}) > 0$, $t = F_i^{\rightarrow}(s_{l_i}^{(k)})$ and recall that the $F_i^{\rightarrow}$ are injective. Now condition (6) yields

$$\|\varphi_l^{(k)}\|_{\mathcal{G}} \leq \Psi\left(\frac{1}{\min\limits_{i=1,\ldots,d} w_{F_i}(s_{i,j_i,l_i}^{(k)} - s_{i,j_i,l_i-1}^{(k)})}\right) \leq \Psi\left(\frac{2^k}{h}\right),$$

since $\min_{i=1,\ldots,d} w_{F_i}(s_{j_i,l_i}^{(k)} - s_{1,j_i,l_i-1}^{(k)}) \geq h2^{-k}$ by (31). $\square$

16

An application of Lemma 4 to (30) yields

$$
\mathrm{P}\Big(\sup_{t\in[t^{(m)}_{j-1},t^{(m)}_j)}|U_n(t)-U^{(m)}_n(t)|\geq\varepsilon\Big)
$$

$$
\leq 2C\Bigg\{\sum_{k=1}^{K_n}\sum_{i=1}^{p}\frac{2^{dk}n^{-(p-i)}}{\varepsilon_k^{2p}}\Big(\frac{3dh}{2^k}\Big)^{\frac{i}{r}}\Phi_i\Big(2\Psi\Big(\frac{2^k}{h}\Big)\Big)
$$

$$
+\sum_{i=1}^{p}\frac{2^{dK_n}n^{-(p-i)}}{(\frac{\varepsilon}{4})^{2p}}\Big(\frac{3dh}{2^{K_n}}\Big)^{\frac{i}{r}}\Phi_i\Big(2\Psi\Big(\frac{2^{K_n}}{h}\Big)\Big)\Bigg\}
$$

$$
\leq D\sum_{i=1}^{p}\Bigg\{n^{-(p-i)}\sum_{k=1}^{K_n}2^{(d-\frac{i}{r})k}k^{4p}\Phi_i\Big(2\Psi\Big(\frac{2^k}{h}\Big)\Big)h^{\frac{i}{r}}\Bigg\}
$$

$$
\leq D\sum_{i=1}^{p-1}\Bigg\{n^{-(p-i)}\Big(\frac{2^{K_n}}{h}\Big)^{d-\frac{i}{r}}K_n^{4p+1}\Phi_i\Big(2\Psi\Big(\frac{2^{K_n}}{h}\Big)\Big)h^{d}\Bigg\}
$$

$$
+D\Bigg\{\sum_{k=1}^{K_n}2^{(d-\frac{p}{r})k}k^{4p}\Phi_p\Big(2\Psi\Big(\frac{2^k}{h}\Big)\Big)h^{\frac{p}{r}}\Bigg\} \tag{32}
$$

for every $j\in\{1,\ldots,m\}^d$, where $D>0$ denotes some finite constant. In the second inequality we used that Ψ and Φ_i are nondecreasing functions and $\varepsilon/4>\varepsilon_{K_n}$.

Let us first deal with the term in the last line of (32). By condition (9) we have

$$
\sum_{k=1}^{K_n}2^{(d-\frac{p}{r})k}k^{4p}\Phi_p\Big(2\Psi\Big(\frac{2^k}{h}\Big)\Big)h^{\frac{p}{r}}\leq C''h^{\frac{p}{r}-\gamma_p}\sum_{k=1}^{\infty}2^{(\gamma_p-(\frac{p}{r}-d))k}k^{4p}
$$

where $\gamma_p<\frac{p}{r}-d$. Hence there is a nonnegative constant $D'<\infty$ such that

$$
\sum_{k=1}^{K_n}2^{(d-\frac{p}{r})k}k^{4p}\Phi_p\Big(2\Psi\Big(\frac{2^k}{h}\Big)\Big)h^{\frac{p}{r}}\leq D'h^{\frac{p}{r}-\gamma_p}=o(h^d). \tag{33}
$$

Now consider the first summand on the right-hand side of inequality (32). In (27) we chose $K_n=\lfloor\log_2(2^4d\sqrt{n}h/\varepsilon)\rfloor$, hence condition (9) yields for any $i=1,\ldots,p-1$,

$$
n^{-(p-i)}\Big(\frac{2^{K_n}}{h}\Big)^{d-\frac{i}{r}}K_n^{4p+1}\Phi_i\Big(2\Psi\Big(\frac{2^{K_n}}{h}\Big)\Big)h^{d}
$$

$$
\leq D''\log_2^{4p+1}\Big(\frac{2^4d}{\varepsilon}\sqrt{n}h\Big)\cdot(\sqrt{n})^{\gamma_i-(\frac{i}{r}+2(p-i)-d)}h^{d}
$$

for some nonnegative constant $D''<\infty$. Since $\gamma_i<\frac{i}{r}+2(p-i)-d$ for $i=1,\ldots,p-1$, by (9) we obtain for all $\eta>0$ and sufficiently large $n\in\mathbb{N}^*$,

$$
D\sum_{i=1}^{p-1}\Bigg\{n^{-(p-i)}\Big(\frac{2^{K_n}}{h}\Big)^{d-\frac{i}{r}}K_n^{4p+1}\Phi_i\Big(2\Psi\Big(\frac{2^{K_n}}{h}\Big)\Big)h^{d}\Bigg\}\leq\frac{1}{2}\eta h^{d}. \tag{34}
$$

Finally, by (32), (33) and (34), for any $\eta>0$

$$
\limsup_{n\to\infty}\mathrm{P}\Big(\sup_{t\in[-\infty,\infty]^d}|U_n(t)-U^{(m)}_n(t)|\geq\varepsilon\Big)
$$

$$
\leq\limsup_{n\to\infty}\sum_{j\in\{1\ldots,m\}^d}\mathrm{P}\Big(\sup_{t\in[t^{(m)}_{j-1},t^{(m)}_j)}|U_n(t)-U^{(m)}_n(t)|\geq\varepsilon\Big)
$$

$$
\leq m^d\Big(o(h^d)+\frac{1}{2}\eta h^d\Big)=m^d\Big(o(m^{-d})+\frac{1}{2}\eta m^{-d}\Big),
$$

since $h = 1/m$. Hence there is a $m_0 \in \mathbb{N}^*$ such that

$$\limsup_{n \to \infty} \mathrm{P}\left(\sup_{t \in [-\infty, \infty]^d} |U_n(t) - U_n^{(m)}(t)| \geq \varepsilon\right) \leq \eta$$

for all $m \geq m_0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

With Lemma 2 and Lemma 3 established, let us finally prove Theorem 1.

Theorem 1. By application of Proposition 4 on $\mathbb{D}([-\infty, \infty]^d)$ equipped with the Skorokhod metric ρ, Lemma 2 (with $r_k^{(m)} := \frac{k}{m}$) and Lemma 3 show that U_n converges in distribution to a process W which is also the limit process of the sequence $W^{(m)}$, $m \in \mathbb{N}^*$. Since all $W^{(m)}$ are centered Gaussian processes the limit process must also be centered Gaussian.

It remains to prove the continuity of the sample paths of W. At this point we already know that U_n converges weakly to W. Therefore it is sufficient to show that for every $\varepsilon, \eta > 0$, there is a $\delta > 0$ such that

$$\limsup_{n \to \infty} \mathrm{P}\left(\sup_{\|t-s\| < \delta} |U_n(t) - U_n(s)| > 3\varepsilon\right) < 3\eta. \tag{35}$$

The sufficiency of this condition can be proven exactly the same way as in the proof of Theorem 15.5 in [2, p.127 f.].

For all $m \in \mathbb{N}^*$, by some triangle inequality arguments we obtain

$$\limsup_{n \to \infty} \mathrm{P}\left(\sup_{\|t-s\| < \delta} |U_n(t) - U_n(s)| > 3\varepsilon\right)$$

$$\leq 2 \limsup_{n \to \infty} \mathrm{P}\left(\sup_t |U_n(t) - U_n^{(m)}(t)| > \varepsilon\right)$$

$$+ \limsup_{n \to \infty} \mathrm{P}\left(\sup_{\|t-s\| < \delta} |U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\right)$$

and thus, by Lemma 3, there is a $m_0 \in \mathbb{N}^*$ such that for all $m \geq m_0$,

$$\limsup_{n \to \infty} \mathrm{P}\left(\sup_{\|t-s\| < \delta} |U_n(t) - U_n(s)| > 3\varepsilon\right)$$

$$\leq 2\eta + \limsup_{n \to \infty} \mathrm{P}\left(\sup_{\|t-s\| < \delta} |U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\right). \tag{36}$$

Now set $\delta_m := \frac{1}{2} \min_{j \in \{0,\ldots,m\}^d} \left\{\max_{i=1,\ldots,d} |t_{j_i} - t_{j_i-1}|\right\}$ and observe that δ_m is strictly positive for any $m \in \mathbb{N}^*$, since the $F_i^{\to}$ used in the construction of the t_j are strictly increasing. Obviously for all $\delta \leq \delta_m$ and $\|t - s\| < \delta$, the points $s, t \in [-\infty, \infty]^d$ must be located in adjacent (or identical) intervals of the form $[t_j, t_{j-\overline{1}})$. Since the process $U_n^{(m)}$ is constant on any of the intervals $[t_j, t_{j-\overline{1}})$ and by symmetry in the arguments s, t we obtain

$$\sup_{\|t-s\| < \delta} |U_n^{(m)}(t) - U_n^{(m)}(s)| = \max_{\substack{j \in \{0,\ldots,m\}^d \\ z \in \{0,1\}^d, \, j \geq z}} |U_n^{(m)}(t_j) - U_n^{(m)}(t_{j-z})|,$$

thus

$$\mathrm{P}\left(\sup_{\|t-s\| < \delta} |U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\right)$$

$$\leq 2^d (m+1)^d \max_{\substack{j \in \{0,\ldots,m\}^d \\ z \in \{0,1\}^d, \, j \geq z}} \mathrm{P}\left(|U_n^{(m)}(t_j) - U_n^{(m)}(t_{j-z})| > \varepsilon\right). \tag{37}$$

Recall that the functions $\varphi_j^{(m)}$ are defined in (19). Analogously to the calculations in Lemma 4, one can show that for all $j \in \{0, \ldots, m\}^d$ and $z \in \{0, 1\}^d$ such that $j \geq z$, we have

$$\|\varphi_{j+\overline{1}}^{(m)}(X_0) - \varphi_{j+\overline{1}-z}^{(m)}(X_0)\|_r \leq \left(\frac{3d}{m}\right)^{\frac{1}{r}},$$
$$\|\varphi_{j+\overline{1}}^{(m)}\|_{\mathcal{G}} \leq \max\left\{\Psi(m), \|1\|_{\mathcal{G}}\right\}.$$

Then, by applying one after another Markov's inequality, the $2p$-th moment bounds (2) and the preceding inequalities, we obtain

$$P\left(|U_n^{(m)}(t_j) - U_n^{(m)}(t_{j-z})| > \varepsilon\right)$$
$$\leq n^{-p}\varepsilon^{-2p}\mathrm{E}\left|\sum_{i=1}^{n}\left(\varphi_{j+\overline{1}}^{(m)}(X_i) - \varphi_{j+\overline{1}-z}^{(m)}(X_i)\right) - \mathrm{E}\left(\varphi_{j+\overline{1}}^{(m)}(X_0) - \varphi_{j+\overline{1}-z}^{(m)}(X_0)\right)\right|^{2p}$$
$$\leq 2Cn^{-p}\varepsilon^{-2p}\sum_{i=1}^{p}n^i\|\varphi_{j+\overline{1}}^{(m)}(X_0) - \varphi_{j+\overline{1}-z}^{(m)}(X_0)\|_r^i \Phi_i\left(2\|\varphi_{j+\overline{1}}^{(m)} - \varphi_{j+\overline{1}-z}^{(m)}\|_{\mathcal{G}}\right)$$
$$\leq 2Cn^{-p}\varepsilon^{-2p}\sum_{i=1}^{p}n^i\left(\frac{3d}{m}\right)^{\frac{i}{r}}\Phi_i\left(2\Psi(m)\right)$$
$$\leq Dm^{\gamma_p - (\frac{p}{r})} + D\sum_{i=1}^{p-1}n^{-(p-i)}m^{\gamma_i - \frac{i}{r}},$$

where D is some finite constant. Therefore by (37) there is another finite constant D' such that

$$P\left(\sup_{\|t-s\|<\delta}|U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\right) \leq D'm^d\left(m^{\gamma_p - \frac{p}{r}} + \sum_{i=1}^{p-1}n^{-(p-i)}m^{\gamma_i - \frac{i}{r}}\right),$$

and thus

$$\limsup_{n\to\infty}P\left(\sup_{\|t-s\|<\delta_m}|U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\right) \leq D'm^{\gamma_p - (\frac{p}{r} - d)} < \eta$$

for sufficiently large $m \in \mathbb{N}^*$, say $m \geq m_1$. By (36) this implies that (35) holds for $\delta = \delta_{\max\{m_0, m_1\}}$. $\qquad\square$

Remark 5. We saw in the proof, that the theorem also holds if (2) is only satisfied for a certain subclass of functions in $\mathcal{G}$, more precisely if (2) holds for all functions $f \in \mathcal{G}$ of the form

$$f := \varphi_{(a,b)} - \varphi_{(a',b')},$$

where $a, b, a', b' \in [-\infty, \infty]^d$, $a' < b$, are such that

$$P(X_0 \in [a', b']) \leq 2P(X_0 \in [a, b]) \leq P(X_0 \in [a', b]) \leq 3P(X_0 \in [a', b']). \tag{38}$$

Choosing for each $m \in \mathbb{N}^*$ an $f_m := \varphi_{(a,b)} - \varphi_{(a',b')}$ such that (38) is satisfied for $P(X_0 \in [a', b']) = 1/m$, it can be shown that

$$\|f_m(X_0)\|_r^i = \mathcal{O}\left(\frac{m^{\gamma_i - \frac{i}{r}}}{\Phi_i(\|f_m\|_{\mathcal{G}})}\right) \quad \text{as } m \to \infty.$$

5 Proof of Proposition 1, Proposition 2 and Proposition 3

5.1 Proof of Proposition 1

By stationarity, we have

$$\left| \mathrm{E}\left(\left(\sum_{i=1}^{n} f(X_i) \right)^p \right) \right| = \left| \sum_{1 \leq i_1,\ldots,i_p \leq n} \mathrm{E}\left(f(X_{i_1}) \cdot \ldots \cdot f(X_{i_p}) \right) \right|$$

$$\leq p! n \left| \sum_{\substack{0 \leq i_1,\ldots,i_{p-1} \leq n-1 \\ i_{p-1}^* \leq n-1}} \mathrm{E}\left(f(X_0) f(X_{i_1^*}) \cdot \ldots \cdot f(X_{i_{p-1}^*}) \right) \right|.$$

Using the notations $I_n(0) := \left| \mathrm{E}(f(X_0)) \right| = 0$ and

$$I_n(p) := \sum_{\substack{0 \leq i_1,\ldots,i_p \leq n-1 \\ i_p^* \leq n-1}} \left| \mathrm{E}\left(f(X_0) f(X_{i_1^*}) \cdot \ldots \cdot f(X_{i_p^*}) \right) \right|, \tag{39}$$

for $p \in \mathbb{N}^*$, we therefore have

$$\left| \mathrm{E}\left(\left(\sum_{i=1}^{n} f(X_i) \right)^p \right) \right| \leq p! n I_n(p-1). \tag{40}$$

Decomposing the sum in (39) with respect to the highest increment of indices i_q, $q \in \{1,\ldots,p\}$, we receive a bound

$$I_n(p) \leq \sum_{q=1}^{p} J_n(p,q),$$

where

$$J_n(p,q) = \sum_{i_q=0}^{n-1} \sum_{\substack{0 \leq i_1,\ldots,i_{q-1},i_{q+1},\ldots,i_p \leq i_q \\ i_p^* \leq n-1}} \left| \mathrm{E}\left(f(X_0) f(X_{i_1^*}) \cdot \ldots \cdot f(X_{i_p^*}) \right) \right|.$$

Lemma 5. *Let $(X_n)_{n \in \mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{G})$, then for all $p \in \mathbb{N}^*$ such that*

$$\sum_{i=0}^{\infty} i^{p-1} \Theta(i) < \infty \tag{41}$$

and all $q \in \{1,\ldots,p\}$ there is a constant K' such that

$$J_n(p,q) \leq K' \|f(X_0)\|_r \|f\|_{\mathcal{G}} + n I_n(q-1) I_n(p-q)$$

for all $n \in \mathbb{N}^*$ and $f \in \mathcal{G}$.

Proof. Set

$$A_{i_1,\ldots,i_p} := \left| \mathrm{Cov}\left(f(X_0) f(X_{i_1^*}) \cdot \ldots \cdot f(X_{i_{q-1}^*}), f(X_{i_q^*}) f(X_{i_{q+1}^*}) \cdot \ldots \cdot f(X_{i_p^*}) \right) \right|$$

$$B_{i_1,\ldots,i_p} := \left| \mathrm{E}\left(f(X_0) f(X_{i_1^*}) \cdot \ldots \cdot f(X_{i_{q-1}^*}) \right) \right| \cdot \left| \mathrm{E}\left(f(X_0) f(X_{i_{q+1}^*}) \cdot \ldots \cdot f(X_{i_p^* - i_q^*}) \right) \right|,$$

where we used the stationarity of $(X_i)_{i \in \mathbb{N}}$ in the last line. We have

$$J_n(p,q)$$
$$\leq \sum_{i_q=0}^{n-1} \sum_{\substack{0 \leq i_1,\ldots,i_{q-1},i_{q+1},\ldots,i_p \leq i_q \\ i_p^* \leq n-1}} A_{i_1,\ldots,i_p} + \sum_{i_q=0}^{n-1} \sum_{\substack{0 \leq i_1,\ldots,i_{q-1},i_{q+1},\ldots,i_p \leq i_q \\ i_p^* \leq n-1}} B_{i_1,\ldots,i_p}.$$

An application of the multiple mixing property (3) yields

$$\sum_{i_q=0}^{n-1} \sum_{\substack{0 \leq i_1,\ldots,i_{q-1},i_{q+1},\ldots,i_p \leq i_q \\ i_p^* \leq n-1}} A_{i_1,\ldots,i_p} \leq K \|f(X_0)\|_r \|f\|_{\mathcal{G}} \sum_{i_q=0}^{n-1} (i_q+1)^{p-1} \Theta(i_q)$$

$$\leq K' \|f(X_0)\|_r \|f\|_{\mathcal{G}}$$

for some constant $K' < \infty$, since $\sum_{i_q=0}^{\infty} i_q^{p-1} \Theta(i_q) < \infty$ by (41). Finally

$$\sum_{i_q=0}^{n-1} \sum_{\substack{0 \leq i_1,\ldots,i_{q-1},i_{q+1},\ldots,i_p \leq i_q \\ i_p^* \leq n-1}} B_{i_1,\ldots,i_p}$$

$$\leq \sum_{i_q=0}^{n-1} \left\{ \sum_{\substack{0 \leq i_1,\ldots,i_{q-1} \leq n-1 \\ i_{q-1}^* \leq n-1}} \left| \mathrm{E}\big(f(X_0) f(X_{i_1^*}) \cdot \ldots \cdot f(X_{i_q^*}) \big) \right| \right.$$

$$\left. \cdot \sum_{\substack{0 \leq i_{q+1},\ldots,i_p \leq n-1 \\ i_p^* - i_q^* \leq n-1}} \left| \mathrm{E}\big(f(X_0) f(X_{i_{q+1}}) \cdot \ldots \cdot f(X_{i_p^* - i_q^*}) \big) \right| \right\}$$

$$= n I_n(p-1) I_n(p-q).$$

$\square$

Lemma 6. *If* $(X_n)_{n \in \mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{G})$, *then for all* $p \in \mathbb{N}^*$ *such that* (41) *is satisfied there is a constant* $K_p < \infty$, *such that*

$$I_n(p) \leq K_p \sum_{i=1}^{\lceil p/2 \rceil} n^{i-1} \|f(X_0)\|_r^i \|f\|_{\mathcal{G}}^i \tag{42}$$

for all $f \in \mathcal{G}$ *with* $\|f\|_{\infty} \leq 1$ *and* $\mathrm{E}(f(X_0)) = 0$.

Proof. We will use mathematical induction to prove the lemma. By Lemma 5 we can easily see that

$$I_n(1) \leq K_1 \|f(X_0)\|_r \|f\|_{\mathcal{G}}$$

for some constant $K_1 < \infty$ if (41) is satisfied. Now consider an arbitrary $\tilde{p} \geq 2$ satisfying (41)

and assume that (42) holds for all $p \leq \tilde{p} - 1$. We have

$$
\begin{aligned}
I_n(\tilde{p}) &\leq \sum_{q=1}^{\tilde{p}} J_n(\tilde{p}, q) \\
&\leq \sum_{q=1}^{\tilde{p}} \left(K' \|f(X_0)\|_r \|f\|_{\mathcal{G}} + n I_n(q-1) I_n(\tilde{p} - q) \right) \\
&\leq K' \tilde{p} \|f(X_0)\|_r \|f\|_{\mathcal{G}} + n \sum_{q=1}^{\tilde{p}} K'' \sum_{i=2}^{\lceil \frac{q-1}{2} \rceil + \lceil \frac{\tilde{p}-q}{2} \rceil} n^{i-2} \|f(X_0)\|_r^i \|f\|_{\mathcal{G}}^i \\
&\leq K_{\tilde{p}} \sum_{i=1}^{\lceil \tilde{p}/2 \rceil} n^{i-1} \|f(X_0)\|_r^i \|f\|_{\mathcal{G}}^i
\end{aligned}
$$

for some constants $K', K'', K_{\tilde{p}} < \infty$, since $\lceil \frac{q-1}{2} \rceil + \lceil \frac{\tilde{p}-q}{2} \rceil \leq \lceil \frac{\tilde{p}}{2} \rceil$. $\qquad \square$

Proposition 1. By (40) and Lemma 6 we immediately obtain

$$
\mathrm{E}\left(\left| \sum_{i=1}^{n} f(X_i) \right|^{2p} \right) \leq (2p)! \, n I_n(2p-1) \leq K_p \sum_{i=1}^{p} n^i \|f(X_0)\|_r^i \|f\|_{\mathcal{G}}^i
$$

since (4) implies that (41) holds with p replaced by $2p - 1$. $\qquad \square$

5.2 Proof of Proposition 2

Since $(X_i)_{i \in \mathbb{N}}$ is a causal function of an i.i.d. process, we have a representation

$$
X_i = G(\xi_i, \xi_{i-1}, \ldots),
$$

with $G : \mathcal{X}^{\mathbb{N}} \to \mathbb{R}^d$. Let $(\xi'_j)_{j \in \mathbb{Z}}$ and $(\xi''_j)_{j \in \mathbb{Z}}$ be copies of the underlying process $(\xi_j)_{j \in \mathbb{Z}}$ such that all three processes are independent. Set

$$
\begin{aligned}
\dot{X}_i^{(k)} &:= G(\xi_i, \xi_{i-1}, \ldots, \xi_{i-k+1}, \xi'_{i-k}, \xi'_{i-k-1}, \ldots) \\
\ddot{X}_i^{(k)} &:= G(\xi_i, \xi_{i-1}, \ldots, \xi_{i-k+1}, \xi''_{i-k}, \xi''_{i-k-1}, \ldots)
\end{aligned}
$$

and note that therefore $(X_i)_{i \in \mathbb{N}} \overset{\mathcal{D}}{=} (\dot{X}_i^{(k)})_{i \in \mathbb{N}} \overset{\mathcal{D}}{=} (\ddot{X}_i^{(k)})_{i \in \mathbb{N}}$. We have

$$
\begin{aligned}
&\left| \mathrm{Cov}\left(f(X_0) \ldots f(X_{i_{q-1}^*}), f(X_{i_q^*}) \ldots f(X_{i_p^*}) \right) \right| \\
&\leq \left| \mathrm{Cov}\left(f(X_0) \ldots f(X_{i_{q-1}^*}) - f(\dot{X}_0^{(k)}) \ldots f(\dot{X}_{i_{q-1}^*}^{(k)}), f(X_{i_q^*}) \ldots f(X_{i_p^*}) \right) \right| \\
&\quad + \left| \mathrm{Cov}\left(f(\dot{X}_0^{(k)}) \ldots f(\dot{X}_{i_{q-1}^*}^{(k)}), f(X_{i_q^*}) \ldots f(X_{i_p^*}) - f(\ddot{X}_{i_q^*}^{(k)}) \ldots f(\ddot{X}_{i_p^*}^{(k)}) \right) \right| \\
&\quad + \left| \mathrm{Cov}\left(f(\dot{X}_0^{(k)}) \ldots f(\dot{X}_{i_{q-1}^*}^{(k)}), f(\ddot{X}_{i_q^*}^{(k)}) \ldots f(\ddot{X}_{i_p^*}^{(k)}) \right) \right|.
\end{aligned} \tag{43}
$$

Since $f(\dot{X}_0^{(k)}) \cdot \ldots \cdot f(\dot{X}_{i_{q-1}^*}^{(k)})$ is $\sigma\left(\{\xi_j : j \leq i_{q-1}^*\} \cup \{\xi'_j : j \in \mathbb{Z}\} \right)$-measurable while $f(\ddot{X}_{i_q^*}^{(k)}) \cdot \ldots \cdot f(\ddot{X}_{i_p^*}^{(k)})$ is $\sigma\left(\{\xi_j : j > i_q^* - k\} \cup \{\xi''_j : j \in \mathbb{Z}\} \right)$-measurable, the functions in the last covariance on the right-hand side of (43) are independent as soon as $k \leq i_q$ and thus the last summand is equal to 0 in this case.

Recall that we only consider such f that satisfy $\|f\|_\infty \le 1$. If we apply Hölder's inequality to equation (43) we obtain for r, s satisfying $\frac{1}{r} + \frac{1}{s} = 1$

$$\left| \mathrm{Cov}\big(f(X_0) \cdot \ldots \cdot f(X_{i_{q-1}^*}) - f(\dot{X}_0^{(k)}) \cdot \ldots \cdot f(\dot{X}_{i_{q-1}^*}^{(k)}), f(X_{i_q^*}) \cdot \ldots \cdot f(X_{i_p^*}) \big) \right|$$

$$\le 2 \| f(X_0) \cdot \ldots \cdot f(X_{i_{q-1}^*}) - f(\dot{X}_0^{(k)}) \cdot \ldots \cdot f(\dot{X}_{i_{q-1}^*}^{(k)}) \|_s \| f(X_{i_q^*}) \cdot \ldots \cdot f(X_{i_p^*}) \|_r$$

$$\le 2q \| f(X_0) \|_r \| f(X_0) - f(\dot{X}_0^{(k)}) \|_s \tag{44}$$

where we used that for $a_i, b_i \in [-1, 1]$,

$$\left| \prod_{i=1}^n a_i - \prod_{i=1}^n b_i \right| \le \sum_{i=1}^n |a_i - b_i|.$$

Since $|f(x) - f(y)| \le \|f\|_{\mathcal{H}_\alpha} \|x - y\|^\alpha$, an application of Jensen's inequality to (44) yields

$$\left| \mathrm{Cov}\big(f(X_0) \cdot \ldots \cdot f(X_{i_{q-1}^*}) - f(\dot{X}_0^{(k)}) \cdot \ldots \cdot f(\dot{X}_{i_{q-1}^*}^{(k)}), f(X_{i_q^*}) \cdot \ldots \cdot f(X_{i_p^*}) \big) \right|$$

$$\le 2q \| f(X_0) \|_r \| f \|_{\mathcal{H}_\alpha} \big(\| X_0 - \dot{X}_0^{(k)} \|_s \big)^\alpha$$

$$= 2q \| f(X_0) \|_r \| f \|_{\mathcal{H}_\alpha} (\delta_{k,s})^\alpha.$$

Analogously we can show that

$$\left| \mathrm{Cov}\big(f(\dot{X}_0^{(k)}) \ldots f(\dot{X}_{i_{q-1}^*}^{(k)}), f(X_{i_q^*}) \ldots f(X_{i_p^*}) - f(\ddot{X}_{i_q^*}^{(k)}) \ldots f(\ddot{X}_{i_p^*}^{(k)}) \big) \right|$$

$$\le 2(p - q) \| f(X_0) \|_r \| f \|_{\mathcal{H}_\alpha} \| (\delta_{k,s})^\alpha,$$

thus for $k = i_q$ we have

$$\left| \mathrm{Cov}\big(f(X_0) \ldots f(X_{i_{q-1}^*}), f(X_{i_q^*}) \ldots f(X_{i_p^*}) \big) \right| \le 2q \| f(X_0) \|_r \| f \|_{\mathcal{H}_\alpha} (\delta_{i_q,s})^\alpha.$$

This proves the first part of the proposition. The second part is a consequence of Proposition 1. □

5.3 Proof of Proposition 3

We will apply the following result obtained by Dedecker in [4].

Proposition 5. *Let* $(Y_i)_{i \in \mathbb{N}}$ *be an ergodic stationary process with* $\mathrm{E}(Y_0) = 0$ *and* $\mathrm{E}(Y_0^2) < \infty$, *which is adapted to a filtration* $(\mathcal{M}_i)_{i \in \mathbb{N}}$. *If*

$$\sum_{i=0}^n Y_0 \mathrm{E}(Y_i | \mathcal{M}_0) \quad \text{converges in } L_1,$$

then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \xrightarrow{\mathcal{D}} N(0, \sigma^2) \quad \text{as } n \to \infty,$$

where $\sigma^2 = \mathrm{E}(Y_0^2) + 2 \sum_{i=1}^\infty \mathrm{E}(Y_0 \cdot Y_i) < +\infty$.

Proposition 3. Choose an arbitrary $f \in \mathcal{H}_\alpha$ with $\mathrm{E}(f(X_0)) = 0$. The process $(Y_i)_{i \in \mathbb{N}}$ given by $Y_i := f(X_i)$ is centered, ergodic, has finite second moments and is adapted to the filtration

$$(\mathcal{M}_i)_{i \in \mathbb{N}} := \big(\sigma(\xi_i, \xi_{i-1}, \ldots) \big)_{i \in \mathbb{N}}.$$

As before, let $(\xi'_j)_{j\in\mathbb{Z}}$ be an independent copy of $(\xi_j)_{j\in\mathbb{Z}}$ and set

$$X'_i := G(\xi'_i, \xi'_{i-1}, \ldots),$$
$$\dot{X}'_i := G(\xi'_i, \xi'_{i-1}, \ldots, \xi'_1, \xi_0, \xi_{-1}, \ldots).$$

Observe that by the independence of $\mathcal{M}_0$ and $\sigma(\{\xi'_i : i \in \mathbb{Z}\})$ we have that

$$\mathrm{E}\big(f(X'_i)|\mathcal{M}_0\big) = \mathrm{E}(f(X'_i)) = 0, \qquad\qquad \mathrm{E}\big(f(X_i)|\mathcal{M}_0\big) = \mathrm{E}\big(f(\dot{X}'_i)|\mathcal{M}_0\big).$$

Thus

$$
\begin{aligned}
\mathrm{E}\Big\{\big|Y_0\mathrm{E}(Y_i|\mathcal{M}_0)\big|\Big\} &\leq \|f\|_\infty \mathrm{E}\Big\{\big|\mathrm{E}(f(X_i)|\mathcal{M}_0)\big|\Big\} \\
&= \|f\|_\infty \mathrm{E}\Big\{\big|\mathrm{E}(f(\dot{X}'_i) - f(X'_i)|\mathcal{M}_0)\big|\Big\} \\
&\leq \|f\|_\infty \mathrm{E}\big|f(\dot{X}'_i) - f(X'_i)\big| \\
&\leq \|f\|_{\mathcal{H}_\alpha}^2 \mathrm{E}\big|\dot{X}'_i - X'_i\big|^\alpha \\
&\leq \|f\|_{\mathcal{H}_\alpha}^2 (\delta_{i,1})^\alpha,
\end{aligned}
$$

where we used Jensen's inequality and $f \in \mathcal{H}_\alpha$ in the last steps. Therefore, by (16),

$$\sum_{i=1}^n Y_0\mathrm{E}\big(Y_i|\mathcal{M}_0\big)$$

converges in L_1 and thus Proposition 5 applies. $\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

References

[1] Berkes, I., Philipp, W.: An almost sure invariance principle for the empirical distribution function of mixing random variables. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **41**(2), 115–137 (1977/78)

[2] Billingsley, P.: Convergence of Probability Measures. John Wiley & Sons Inc., New York (1968)

[3] Borovkova, S., Burton, R., Dehling, H.: Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation. Trans. Amer. Math. Soc. **353**(11), 4261–4318 (electronic) (2001)

[4] Dedecker, J.: A central limit theorem for stationary random fields. Probab. Theory Related Fields **110**(3), 397–426 (1998)

[5] Dedecker, J.: An empirical central limit theorem for intermittent maps. Probab. Theory Related Fields **148**(1-2), 177–195 (2010)

[6] Dedecker, J., Prieur, C.: An empirical central limit theorem for dependent sequences. Stochastic Process. Appl. **117**(1), 121–142 (2007)

[7] Dehling, H., Durieu, O.: Empirical processes of multidimensional systems with multiple mixing properties. Stochastic Process. Appl. **121**(5), 1076–1096 (2011)

[8] Dehling, H., Durieu, O., Volny, D.: New techniques for empirical processes of dependent data. Stochastic Process. Appl. **119**(10), 3699–3718 (2009)

[9] Dehling, H., Philipp, W.: Empirical process techniques for dependent data. In: Empirical Process Techniques for Dependent Data, pp. 3–113. Birkhäuser Boston, Boston, MA (2002)

[10] Donsker, M.D.: Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. Ann. Math. Statistics **23**, 277–281 (1952)

[11] Doukhan, P., Louhichi, S.: A new weak dependence condition and applications to moment inequalities. Stochastic Process. Appl. **84**(2), 313–342 (1999)

[12] Doukhan, P., Massart, P., Rio, E.: Invariance principles for absolutely regular empirical processes. Ann. Inst. H. Poincaré Probab. Statist. **31**(2), 393–427 (1995)

[13] Hennion, H., Hervé, L.: Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness, *Lecture Notes in Mathematics*, vol. 1766. Springer-Verlag, Berlin (2001)

[14] Le Borgne, S.: Limit theorems for non-hyperbolic automorphisms of the torus. Israel J. Math. **109**, 61–73 (1999)

[15] Leonov, V.P.: On the central limit theorem for ergodic endomorphisms of compact commutative groups. Dokl. Akad. Nauk SSSR **135**, 258–261 (1960)

[16] Neuhaus, G.: On weak convergence of stochastic processes with multidimensional time parameter. Ann. Math. Statist. **42**, 1285–1295 (1971)

[17] Straf, M.L.: Weak convergence of stochastic processes with several parameters. In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability Theory, pp. 187–221. Univ. California Press, Berkeley, Calif. (1972)

[18] Wu, W.B.: Nonlinear system theory: another look at dependence. Proc. Natl. Acad. Sci. USA **102**(40), 14150–14154 (2005)

[19] Wu, W.B.: Empirical processes of stationary sequences. Statist. Sinica **18**(1), 313–333 (2008)